DETECTION AND ESTIMATION OF SPIKES IN PRESENCE OF NOISE AND INTERFERENCE

Damien Passemier, Abla Kammoun, Merouane Debbah

DETECTION AND ESTIMATION OF SPIKES IN PRESENCE OF NOISE AND INTERFERENCE

Damien Passemier[2] and Abla Kammoun[1] and Mérouane Debbah[1]

Alcatel-Lucent Chair, Supélec[1], Hong Kong University of Science and Technology [2]

ABSTRACT

In many practical situations, the useful signal is contained in a low-dimensional subspace, drown in noise and interference. Many questions related to the estimation and detection of the useful signal arise. Because of their particular structure, these issues are in connection to the problem that the mathematics community refers to as ”spike detection and estimation”. Previous works in this direction have been restricted to either determining the number of spikes or estimating their values while knowing their multiplicities. This motivates our work which considers the joint estimation of the number of spikes and their corresponding orders, a problem which has not been yet investigated to the best of our knowledge.

1. INTRODUCTION

Detecting and estimating the components of a signal corrupted by additive Gaussion noise is a fundamental problem that arises in many signal and array processing applications. Considering a large number of received samples, one can easy see that their covariance matrix exhibit a different behaviour depending on the number of the components of the useful signal. In light of this consideration, first methods of signal detection like techniques using the Roy Test [1] or those using information theoretic criteria [2] have been based on the eigenvalues of the empirical covariance matrix. Recently, the advances in the spectral analysis of large dimensional random matrices have engendered a new wave of interest for the scenario when the number of observations is of the same order of magnitude as the dimension of the received samples, while the number of signal components remain finite. Such a model is referred to as the spiked covariance model [6]. It has allowed the emergence of new detection schemes based on the works of the extreme eigenvalues of large random Wishart matrices [3, 4, 5]. It is especially encountered in multi-sensor detection [7] and power estimation problems [8], which are at the heart of cognitive radio applications. This model has also found application in subspace estimation problems with a particular interest on the estimation of directions of arrival [13].

This research has been supported by the ERC Starting Grant 305123 MORE (Advanced Mathematical Tools for Complex Network Engineering).

From a mathematical perspective, the focus has been either to detect the presence of sources and estimate their numbers [10, 11] or to estimate their powers [12]. The general case where the objective is to extract as much as possible information has not been addressed to the best of our knowledge. This motivates our work which proposes an easy way to jointly estimate the number of sources, their powers and their multiplicities in the case where different sources are using the same power values.

2. SYSTEM MODEL

Consider the observation vector $\mathbf{x}_i \in \mathbb{C}^p$ at time i:

$$\mathbf{x}_i = \sum_{k=1}^{K} \sqrt{\alpha_k}\mathbf{W}_k\mathbf{s}_{k,i} + \sigma\mathbf{n}_i$$

where

- $(\mathbf{W}_k)_{k=1,\dots,m}$ is an orthogonal family of rectangular unitary $p \times m_k$ matrices (i.e, $\forall k, \mathbf{W}_k$ has orthogonal columns and $\mathbf{W}_k\mathbf{W}_j^{\mathrm{H}} = 1_{k=j}\mathbf{I}_p$);
- α_k are positive distinct scalars: $\alpha_1 > \alpha_2 > \cdots > \alpha_K$;
- $\mathbf{s}_{k,i} \in \mathbb{C}^{m_k} \times 1$ are independent random vectors with mean zero and variance 1;
- $\mathbf{n}_i \in \mathbb{C}^{p\times 1}$ is complex Gaussian distributed (i.e. $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$) and represent the interference and noise signal.

Gathering n observations $\mathbf{x}_1, \dots, \mathbf{x}_n$ into a $p \times n$ observation matrix $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$, we obtain

$$\mathbf{X} = [\mathbf{W}_1, \cdots, \mathbf{W}_K] \begin{bmatrix} \sqrt{\alpha_1}\mathbf{I}_{m_1} & \cdots & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\alpha_K}\mathbf{I}_{m_K} \end{bmatrix} \times \begin{bmatrix} \mathbf{s}_{1,1} & \cdots & \mathbf{s}_{1,n} \\ \vdots & & \vdots \\ \mathbf{s}_{K,1} & \cdots & \mathbf{s}_{K,n} \end{bmatrix} + \sigma\left[\mathbf{n}_1, \cdots, \mathbf{n}_n\right].$$

or equivalently

$$\mathbf{X} = \mathbf{\Sigma}^{\frac{1}{2}}\mathbf{Y} \tag{1}$$

where $\mathbf{Y}$ is a matrix of independent entries with zero mean and variance 1 and $\mathbf{\Sigma}$ is the theoretical covariance matrix of the observations given by:

$$\mathbf{\Sigma} = \begin{bmatrix} (\alpha_1+\sigma^2)\mathbf{I}_{m_1} & \mathbf{0} & \cdots & \mathbf{0} \\ & \ddots & & \\ & & (\alpha_K+\sigma^2)\mathbf{I}_{m_K} & \\ & & & \sigma^2\mathbf{I}_{n-m} \end{bmatrix}$$

where $m = \sum_{k=1}^{K} m_k$. Note that $\mathbf{\Sigma}$ has K eigenvalues with multiplicities $m_1, \ldots, m_K$ and one eigenvalue equal to σ^2 with multiplicity $p - m$, where .

This model is a generalization of the spiked covariance model, since the covariance matrix $\mathbf{\Sigma}$ is allowed to have multiple eigenvalues. It can be encountered as shown in [8] for power estimation purposes in cognitive radio networks. Another interesting application is met in the array processing field and in particular in the problem of the estimation of the angles of arrival. In this case, the received signal matrix is given by [13]:

$$\mathbf{X} = \mathbf{A}(\theta)\mathbf{P}^{\frac{1}{2}}\mathbf{S} + \sigma\mathbf{N} \tag{2}$$

where $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_m)]$, $\mathbf{a}(\theta_i)$ being the steering vector, $\mathbf{P} = \mathrm{diag}\left(\alpha\mathbf{I}_{m_1}, \ldots, \alpha_K\mathbf{I}_{m_K}\right)$ and $\mathbf{S}$ is the $m \times n$ transmitted matrix of i.i.d Gaussian entries. Note that in this case, $\mathbf{A}$ can be considered as unitary, since: $\mathbf{A}^{\mathsf{H}}\mathbf{A} \xrightarrow[p\to+\infty]{} \mathbf{I}_m$.

Previous methods dealing with the estimation of directions of arrivals has so far assumed a prior estimation of the number of sources [13]. Such information is obviously not always available in practice. This motivates our paper, which proposes a method to jointly estimate the number of sources as well as their multiplicities.

3. ESTIMATION OF SPIKES' VALUES AND MULTIPLICITIES

The estimation technique relies on results about the asymptotic behavior of the covariance matrix. As shown in the following theorem proven in [15], the asymptotic spectral properties of the covariance matrix depend on the eigenvalues $\alpha_1, \cdots, \alpha_K$ of matrix $\mathbf{\Sigma}$.

Theorem 1. *Let $\mathbf{S}_n$ be the sample covariance matrix given by:*

$$\mathbf{S}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{x}_i\mathbf{x}_i^{H}$$

Denote by $\widehat{\lambda}_{n,1} > \widehat{\lambda}_{n,2} > \cdots > \widehat{\lambda}_{n,p}$ the p eigenvalues of $\mathbf{S}_n$ arranged in decreasing order. Let $s_i = \sum_{k=1}^{i} m_k$ and J_k the index set $J_k = \{s_k+1, \ldots, s_k+m_k\}$, $k \in \{1, \ldots, K\}$.

Assume that $\gamma_n = \frac{p}{n} \to \gamma$, and let $\phi(x) = x + \sigma^2 + \gamma\sigma^2\left(1+\frac{\sigma^2}{x}\right)$ for $x \neq 0$. Then, for any $k \in \{1, \ldots, K\}$, if $\phi'(\alpha_k) > 0$, i.e, $\alpha_K > \sigma^2\sqrt{\gamma}$, we have,

$$\widehat{\lambda}_{n,j} \to \phi(\alpha_k), \ \ \forall j \in J_k$$

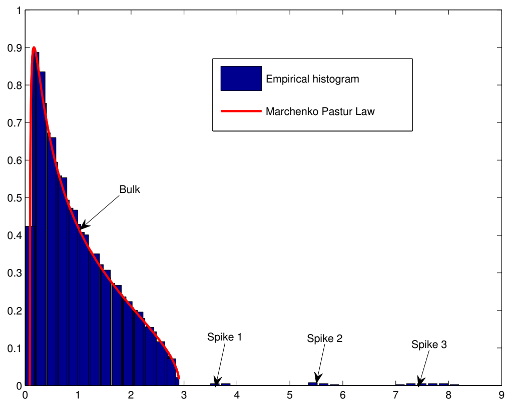

Fig. 1. Histogram of the eigenvalues of the empirical covariance matrix.

Remark 1. *Under the condition $\phi'(\alpha_k) > 0$ for all $k \in \{1, \ldots, K\}$, the empirical distribution of the spectrum is composed of $K+1$ connected intervals: a bulk corresponding to the Marčhenko-Pastur law [14] followed by K spikes. To illustrate this, we represent in Fig. 1, the empirical histogram of the eigenvalues of the empirical covariance matrix when $K = 3$, $(\alpha_1, \alpha_2, \alpha_3) = (7, 5, 3)$, $(n, p) = (4000, 2000)$ and $\sigma^2 = 1$.*

Fig. 1 provides us insights about an intuitive approach to estimate the multiplicities of spikes and their values given their number K. Actually, one needs to rearrange the eigenvalues and then detect the largest gaps that correspond to a switch from one connected interval to the next one.

This leads us to distinguish two cases whether K is either known or not. We will consider these cases in the following:

3.1. K is known

In this case, we propose to estimate the eigenvalues by considering the differences between consecutive eigenvalues:

$$\delta_{n,j} = \widehat{\lambda}_{n,j} - \widehat{\lambda}_{n,j+1}, \ \ j \geq 1.$$

Indeed, the results quoted above imply that a.s. $\delta_{n,j} \to 0$, for $j \notin \{s_i, i = 1, \ldots, K\}$ whereas for $j \in \{s_i, i = 1, \ldots, K\}$, $\delta_{n,j}$ tends to a positive limit given by $\phi(\alpha_j) - \phi(\alpha_{j+1})$. Thus it becomes possible to estimate the multiplicities from index-numbers j where $\delta_{n,j}$ is large. If K is known, we will take the indices corresponding to the K larger differences $\delta_{n,i}$. Denote by $i_1, \ldots, i_p$ the indices of the differences $\delta_{n,i}$ such that $\delta_{n,i_1} \geq \cdots \geq \delta_{n,i_p}$. Then, the estimator $(\hat{m}_1, \ldots, \hat{m}_K)$ of the multiplicities $(m_1, \ldots, m_K)$ is defined by

$$\begin{cases} \hat{m}_1 = \min\{i_k, k \in \{1, \ldots, K\}\} \\ \hat{m}_2 = \min\{i_k, k \in \{1, \ldots, K\} \backslash \{\hat{m}_1\}\} - \hat{m}_1 \\ \hat{m}_j = \min\{i_k, k \in \{1, \ldots, K\} \backslash \{\hat{m}_1, \ldots, \hat{m}_{j-1}\}\} - \sum_{i=1}^{j-1}\hat{m}_i \\ \hat{m}_K = \max\{i_k, k \in \{1, \ldots, K\}\} - \sum_{i=1}^{K-1}\hat{m}_i \end{cases}$$

The proposed consistent estimator of the number of the spikes is therefore given by the following theorem, for which a proof is omitted because of lack of space:

Theorem 2. *Let* $(\mathsf{x}_i)_{1\leq i\leq n}$ *be* n *copies i.i.d. of* x *which follows the model* (1)*. Suppose that the population covariance matrix* Σ *has* K *non null and non unit eigenvalues* $\alpha_1 > \cdots > \alpha_K > \sigma^2\sqrt{\gamma}$ *with respective multiplicity* $(m_k)_{1\leq k\leq K}$ *(*$m_1 + \cdots + m_K = m$*), and* $p - m$ *eigenvalues equal to* σ^2*. Assume that* $\frac{p}{n} \to \gamma > 0$ *when* $n \to \infty$*. Then the estimator (*$\hat{m}_1, \ldots, \hat{m}_K$*) is strongly consistent, i.e* $(\hat{m}_1, \ldots, \hat{m}_K) \to (m_1, \ldots, m_K)$ *almost surely when* $n \to \infty$*.*

3.2. K is not known

As fig. 1 shows, eigenvalues outside the bulk are organized into K clusters, where within each cluster, all eigenvalues converge in the asymptotic regime $p, n \to +\infty, p/n \to \gamma$ to the same value. If K is not estimated correctly, applying the previous method, will lead to either gathering two close clusters (K is under-estimated) or to subdividing the clusters corresponding to the highest spikes (K is over-estimated). Clearly, the second order results within each cluster seems to bring useful information which allows to discard these cases. In particular, in the sequel, we will rely on the following theorem which is a by-product of Proposition 3.2 in [15]:

Theorem 3. *Assume that the setting of Theorem 2 holds. Let* $g_k = \sum_{j=s_{k-1}+1}^{s_k} \widehat{\lambda}_{n,j}$*, the sum of the eigenvalues corresponding to the* k*-th cluster. Then* g_k *verify*

$$\sqrt{n}\left(\sum_{j=s_{k-1}+1}^{s_k} \widehat{\lambda}_{n,j} - m_k\phi(\alpha_k)\right) \xrightarrow[n,p\to\infty]{\mathcal{L}} \mathcal{N}(0, 2m_k v_k^2)$$

where $v_k^2 = \frac{2\alpha_k'^2((\alpha_k'-1)^2-\gamma)}{(\alpha_k'-1)^2}$, $\alpha_k' = \frac{\alpha_k}{\sigma^2} + 1$ *and* $\mathcal{L}$ *denotes the convergence in distribution.*

Theorem 3 establishes that the sum of the eigenvalues within the k-th cluster behaves as a Gaussian random variable with mean and variance depending on the unknown value α_k. One way to remove the uncertainty in the unknowns α_k is to assume that they are random with a priori known distribution $\pi(\alpha_1, \ldots, \alpha_K|K)$. A possible case would correspond to the situation where they are uniformly distributed over a finite discrete set[1].

Since the clusters are asymptotically independent [17], the likelihood function (distribution of $\mathbf{g} = [g_1, \cdots, g_K]$ under the underlying parameters $\alpha_1, \cdots, \alpha_K, m_1, \cdots, m_K, K$ is given by:

$$f(\mathbf{g}|\alpha_1, \ldots, \alpha_K, K) = \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi v_k^2}} e^{-\frac{1}{2v_k^2}(g_k - m_k\phi\alpha_k)^2}$$

[1] A discrete distribution for powers has been considered in [16].

where the multiplicities $m_1, \cdots, m_K$ can be estimated in a consistent way given the number of classes K as it has been shown in section 3.1. Hence, the maximum likelihood function $f(\mathbf{g}|K)$ is given by:

$$f(\mathbf{g}|K) = \mathbb{E}\left[f(\mathbf{g}|\alpha_1, \ldots, \alpha_K, K)\right] \quad (3)$$

where the expectation is taken over the a priori distribution $\pi(\alpha_1, \cdots, \alpha_K|K)$. The maximum likelihood estimator $\widehat{K}$ is thus given by:

$$\widehat{K} = \max_{1\leq k\leq m} \mathbb{E}\left[f(\mathbf{g}|\alpha_1, \ldots, \alpha_K, K)\right].$$

Once K is estimated, the multiplicities can be retrieved by using the method in Section 3.1.

To sum up, when K is unknown, the estimation of the unknown parameters using the a priori π consists in the following steps :

1. Compute the consecutive differences $\delta_{n,j} = \widehat{\lambda}_{n,j} - \widehat{\lambda}_{n,j+1}$ of the ordered eigenvalues of the sample covariance matrix $\mathbf{S}_n$;
2. For each k ranging from one to $K_{\max}$, (where $K_{\max}$ is a known upper bound of K) calculate the corresponding estimator of the multiplicities $(\hat{m}_1^{(k)}, \ldots, \hat{m}_k^{(k)})$ using Theorem 2, and compute the maximum likelihood function (3).
3. Select K such that it maximizes the maximum likelihood function.

4. NUMERICAL EXPERIMENTS

We consider in our simulations the model described by (2) given in section 2 with $\mathbf{A}(\theta) = p^{-1/2}\left[\exp\left(-iv\sin(\theta)\pi\right)\right]_{v=0}^{p-1}$. We assume that the set of the a priori spikes is $E = \{1, 3, 5, 7\}$ and that the values $\alpha_1, \ldots, \alpha_K$ are uniformly distributed over this set.

In the sequel, we will display the empirical probability $\mathbb{P}(\hat{K} = K)$ calculated over 500 independent realizations. For each iteration, we choose the "true" values of spikes uniformly in the set E, but with the same fixed proportion m_i/m, $i = 1, \ldots, K$.

We consider two different experiments: in the first one, we study the performance of our method for different level of noise variances whereas for the second one, we consider the impact of the number of spikes m for a fixed noise variance.

4.1. Performance of the proposed method with respect to the SNR

In this experiment, we consider the detection of the number of $K = 3$ different clusters of 500×1 ($p = 500$) signals from $n = 1000$ samples. We assume that the unknown multiplicities are $m_1 = 1$, $m_2 = 4$, $m_3 = 2$. Since the minimum

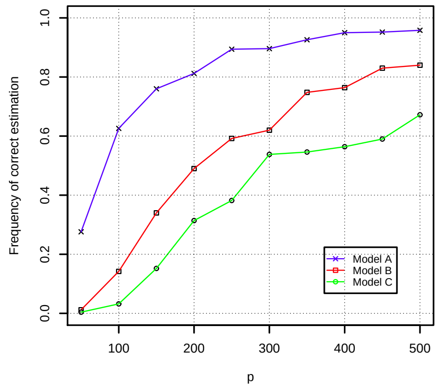

Fig. 2. Empirical probability of $\mathbb{P}(\hat{K} = K)$ as a function of (p, n), for Models A, B and C.

value of the spike is assumed to be 1, σ^2 has to be lower than $1/\sqrt{c} = 1.4142$ in order to keep a gap between $\hat{\lambda}_m$ and $\hat{\lambda}_{m+1}$ (see Theorem 1). The noise variance is expressed in dB $10\log_{10}(\sigma^2)$. Table 1 illustrates the obtained results :

Table 1. Empirical probability of $\mathbb{P}(\hat{K} = K)$ as a function of the σ^2.

σ^2(dB)	-50	-40	-30	-20	-10	-6.99	-5.223	-3.98	-3.01	-2.22
$\mathbb{P}(\hat{K} = K)$	0.992	0.978	0.988	0.986	0.984	0.978	0.978	0.980	0.964	0.974
SNR (dB)	-1.55	-0.97	-0.46	0	0.41	0.80	0.97	1.14	1.30	1.46
$\mathbb{P}(\hat{K} = K)$	0.972	0.954	0.960	0.968	0.942	0.926	0.896	0.850	0.694	0.476

Our estimator performs well, especially for low noise variances. When σ^2 is getting close to the threshold 1.41 (i.e. 1.50 dB), the estimator becomes less accurate, which was expected since $\hat{\lambda}_m$ is very close to the bulk.

4.2. Influence of the number of spikes m

We study in this experiment the impact of the number of spikes in the performance of the proposed estimation method. Similarly to the previous simulation setting, we set $K = 3$ and $\gamma = \frac{p}{n} = 0.5$.

Figure. 2 displays the frequency of correct estimation of the following three models with respect to p:

- Model A: $m = 4$, with $m_1 = 1$, $m_2 = 2$, $m_3 = 1$;
- Model B: $m = 8$, with $m_1 = 2$, $m_2 = 4$, $m_3 = 2$;
- Model C: $m = 12$, with $m_1 = 3$, $m_2 = 6$, $m_3 = 3$;

Note that these three models keep the $\frac{p}{n}$ and $\frac{m_i}{m}$ fixed except $\frac{m}{n}$ which is different. In that way, only the impact of the variation of the number of spikes is visualized.

As expected, our estimator performs better in Model A than in Model B and C. In both cases, we observe the asymptotic consistency, but the convergence is slower for Model C.

Remark 2. *We have noticed by simulations, once K was correctly estimated, the multiplicities are correctly estimated, an observation which is in accordance with our Theorem 2.*

5. CONCLUSION

The problem of signal detection appears naturally in many signal processing applications. Previous works used to deal with this problem partially by assuming extra knowledge about the number of spikes or their corresponding orders. This work is therefore an attempt to consider the general problem where the objective is to estimate all the unknown parameters. In particular, we show that when the number of different spikes is known, their multiplicities can be estimated consistently. In light of this consideration, we propose a bayesian estimation method which jointly infer the number of spikes and their multiplicities. The experiments that we carried out support the performance of the proposed technique

6. REFERENCES

[1] S. N. Roy, *Some Aspects of Mutlivariate Analysis*, Wiley, New-York, 1957.

[2] M. Wax and T. Kailath, “Detection of Signals by Information Theoretic Criteria,” *IEEE Trans. Acoust. Speech. Signal Process.*, vol. 33, no. 2, pp. 387–392, 1985.

[3] B. Nadler, “Nonparametric Detection of Signals by Information Theoretic Criteria : Performance Analysis and an Improved Estimator,” *IEEE Transactions on Signal Processing*, vol. 58, no. 5, 2010.

[4] B. Nadler, “Detection Performance of Roy’s Largest Root Test When the Noise Covariance Matrix is Arbitrary,” in *ssp*, Nice, 2011.

[5] R. R. Nadakuditi and J. W. Silverstein, “Fundamental Limit of Sample Generalized Eigenvalue Based Detection of Signals in Noise Using Relatively few Signal-bearing and noise-only samples,” *IEEE Selected Topics in Signal Processing*, vol. 4, no. 3, pp. 468–480, 2010.

[6] I. M. Johnstone, “On the Distribution of the Largest Eigenvalue in Principal Component Analysis,” *Annals of Statistics*, vol. 29, no. 2, pp. 295–327, 2001.

[7] Federico Penna, *Statistical Methods for Cooperative and Distributed Inference in Wireless Networks*, Ph.D. thesis, Politecnico di Torino, Torino, Italy, Mar. 2012.

[8] J. Yao, R. Couillet, J. Najim, E. Moulines, and M. Debbah, “CLT for eigen-inference methods for cognitive radio,” in *ICASSP*, Prague, 2011.

[9] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet, ,” *Journal of Multivariate Analysis*, vol. 114, pp. 427–447, Feb. 2013.

[10] S. Kritchman and B. Nadler, “Non-Parametric Detection of the Number of Signals: Hypothesis Testing and Random Matrix Theory,” *IEEE Transactions on Signal Processing*, vol. 57, no. 10, Oct. 2009.

[11] D. Passemier and J. F. Yao, “Estimation of the Number of Factors, Possibly Equal, in the High Dimension Case,” *Submitted to Journal Of Multivariate Analysis*, 2011.

[12] Z. Bai and X. Ding, “Estimation of Spiked Eigenvalues in Spiked Models,” *Random Matrices: Theory and Applications*, vol. 1, no. 2, 2012.

[13] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet, “A Subspace Estimator for Fixed Rank Perturbations of Large Random Matrices,” *Journal of Multivariate Analysis*, pp. 427–447, 2013.

[14] J. Baik and J. W. Silverstein, “Eigenvalues of Large Sample Covariance Matrices of Spiked Population Models,” *Journal of Multivariate Statistics*, vol. 97, pp. 1382–1408, 2006.

[15] Z. D. Bai and J. F. Yao, “Central Limit Theorems for Eigenvalues in a Spiked Population Model,” *Ann. Inst, Henri Poincaré*, pp. 447–474, 2008.

[16] J.-M. Chaufray, W. Hachem, , and P. Loubaton, “Asymptotic analysis of optimum and sub-optimum CDMA downlink MMSE receivers,” *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2620–2638, 2004.

[17] R. Couillet and W. Hachem, “Fluctuations of Spiked Random Matrix Models and Failure Diagnosis in Sensor Networks,” *IEEE Transactions on Information Theory*, pp. 509–525, Jan. 2013.